



Joint classification of multiresolution and multisensor data using a multiscale Markov mesh model

Alessandro Montaldo, Luca Fronda, Ihsen Hedhli, Gabriele Moser, Josiane Zerubia, Sebastiano B. Serpico

► To cite this version:

Alessandro Montaldo, Luca Fronda, Ihsen Hedhli, Gabriele Moser, Josiane Zerubia, et al.. Joint classification of multiresolution and multisensor data using a multiscale Markov mesh model. IGARSS 2019 - IEEE International Geoscience and Remote Sensing Symposium, Jul 2019, Yokohama, Japan. hal-02157082

HAL Id: hal-02157082

<https://inria.hal.science/hal-02157082>

Submitted on 15 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

JOINT CLASSIFICATION OF MULTIREOLUTION AND MULTISENSOR DATA USING A MULTISCALE MARKOV MESH MODEL

*Alessandro Montaldo¹, Luca Fronda¹, Ihsen Hedhli³,
Gabriele Moser¹, Josiane Zerubia², Sebastiano B. Serpico¹*

¹ University of Genoa, DITEN dept., Genoa, Italy. gabriele.moser@unige.it

² Inria, UCA, Sophia-Antipolis Méditerranée Center, France. ³ Laval University, Quebec, Canada

ABSTRACT

In this paper, the problem of the classification of multiresolution and multisensor remotely sensed data is addressed by proposing a multiscale Markov mesh model. Multiresolution and multisensor fusion are jointly achieved through an explicitly hierarchical probabilistic graphical classifier, which uses a quadtree structure to model the interactions across different spatial resolutions, and a symmetric Markov mesh random field to deal with contextual information at each scale and favor applicability to very high resolution imagery. Differently from previous hierarchical Markovian approaches, here, data collected by distinct sensors are fused through either the graph topology itself (across its layers) or decision tree ensemble methods (within each layer). The proposed model allows taking benefit of strong analytical properties, most remarkably causality, which make it possible to apply time-efficient non-iterative inference algorithms.

Index Terms— Multiresolution and multisensor fusion, symmetric Markov mesh, hierarchical MRF, tree ensemble

1. INTRODUCTION

In this paper we address the problem of data fusion for classification purposes, in the context of different types of very high resolution (VHR) remotely sensed images by proposing a hierarchical hidden Markov model. The proposed methodology, which partly extends the previous methods in [1] and [2], aims at addressing the challenging task of jointly classifying data taken by different sensors at different spatial resolutions, while maintaining an appropriate trade-off between accuracy and computation time. On one hand, this is a typical scenario from an application-oriented perspective, owing to the potential offered by the current space missions with optical (e.g., Pléiades, WorldView-3, SPOT-6/7) and synthetic aperture radar (SAR; e.g., COSMO-SkyMed, TerraSAR-X, RADARSAT-2) VHR payloads in terms of data acquisition capabilities. On the other hand, it is a scarcely explored and difficult problem because of the substantially heterogeneous statistics of the data collected by different sensors (e.g., optical and SAR) and of the need to capture the different spatial

details associated with their resolutions [1]. In this paper, we leverage on the modeling power of causal probabilistic graphical models to address this task.

In particular, Markov random field (MRF) models on planar graphs have been widely used for classification since they provide a consistent framework for integrating spatial-contextual information into Bayesian classification. However, because of their generally non-causal nature, they often lead to iterative and computationally intensive inference algorithms [3]. On the contrary, hierarchical MRFs, whose topology is based on multiscale graphs such as quadrees, exhibit causality properties that make highly efficient inference procedures feasible [4]. MRFs based on systems of quadrees have also been found effective in [1] to jointly classify multisensor and multiresolution remote sensing data. However, a common limitation of a hierarchical MRF on quadtree is that, while it captures relations among data at different scales through a Markov chain, it generally does not explicitly characterize spatial dependences within the layer at each scale [4].

In this paper, a hierarchical Markov mesh random field (MMRF) is proposed to classify multiresolution and multisensor remote sensing data, leveraging on the aforementioned benefits of both hierarchical and planar Markov models. In a classical MMRF model, a planar lattice is endowed with a case-specific neighborhood system that allows causality to be formulated [5, 6]. Here, the model introduced in [2] is adopted, in which hierarchical Markovianity with respect to a quadtree is postulated, while an MMRF is defined on each layer of the quadtree. This combination ensures causality for the whole probabilistic graphical model – thus favoring efficient inference algorithms –, takes benefit from the spatial-contextual information within each layer, and intrinsically supports multiresolution fusion.

In [2], this model was formulated in combination with Gaussian mixtures to consider the specific framework of the fusion of panchromatic and multispectral images of urban areas. Here, we generalize the model to the broader case of multisensor data involving both optical and SAR components. The data coming from the input sensors are inserted in the

quadtree based on their resolutions. A first novel contribution of this work is the formalization of the fusion of data collected by sensors operating at different resolutions through the inference process across the hierarchical topology. A second novel element is the integration of nonparametric ensemble modeling into the hierarchical MMRF to also allow for the fusion of data taken by multiple sensors operating at the same resolution. Various decision tree ensembles [7] are used for this purpose. In particular, a common shortcoming of an MMRF is that its spatial formulation of causality may lead to undesired directional artifacts (“corner dependence”) [5, 6]. To remove this drawback, the approach in [6] is adopted, in which a symmetric, corner-independent, isotropic formulation is defined and incorporated within the proposed hierarchical model.

2. METHODOLOGY

2.1. Hierarchical Markov mesh random field

First, let $S \subset \mathbb{Z}^2$ be a rectangular pixel lattice, each site $s \in S$ being associated with a discrete random variable x_s . $\mathcal{X} = \{x_s\}_{s \in S}$ is the corresponding random field. Let us also define an order relation $<$ on S , so that, intuitively, it makes sense to speak of the “past” of each site $s \in S$ (i.e., of all sites $r \in S$ such that $r < s$). A neighborhood relation is also assumed to be defined in S consistently with the order relation, and we write $r \lesssim s$ to indicate that r is a causal neighbor of s . \mathcal{X} is an MMRF if, for all $s \in S$, [5]:

$$P(x_s | x_r, r < s) = P(x_s | x_r, r \lesssim s), \quad (1)$$

i.e., if Markovianity holds when restricted to the past of each site. A typical choice is that the past of s is the set of all pixels reached by a raster scan before getting to s and that the neighbors of s are the three nearby pixels located in its previous row and column (second-order causal neighborhood). An MMRF is an MRF [5], whereas an arbitrary MRF is generally not an MMRF because it lacks causality.

Then, let $\{S^0, S^1, \dots, S^L\}$ be a collection of pixel lattices organized as a quadtree, i.e., let the width and height of $S^{\ell-1}$ be twice smaller than those of S^ℓ , so that each site $s \in S^\ell$ has a well-defined parent site $s^- \in S^{\ell-1}$ and four well-defined children sites in $S^{\ell+1}$ (collected in a set $s^+ \subset S^{\ell+1}$; $\ell = 1, 2, \dots, L-1$). This structure implies a hierarchy on the whole tree $S = \bigcup_{\ell=0}^L S^\ell$ from the root to the leaves [4]. If a discrete random variable x_s is associated with each $s \in S$, then $\mathcal{X} = \{x_s\}_{s \in S}$ is a hierarchical MRF if [4]:

$$P(\mathcal{X}^\ell | \mathcal{X}^{\ell-1}, \mathcal{X}^{\ell-2}, \dots, \mathcal{X}^0) = P(\mathcal{X}^\ell | \mathcal{X}^{\ell-1}), \quad (2)$$

where $\mathcal{X}^\ell = \{x_s\}_{s \in S^\ell}$ ($\ell = 1, 2, \dots, L$), i.e., if Markovianity holds across the scales.

More details on mesh and hierarchical MRFs can be found in [1, 2, 4, 5]. Here, both (1) and (2) are assumed to hold jointly, thus defining a hierarchical MMRF. The rationale is twofold. First, the hierarchical model naturally

fits the requirements of multiresolution fusion. Second, the mesh model allows spatial relations to be captured within each single-resolution layer as well, a desirable property with VHR images. In the hierarchical MMRF, each node in each layer except the root is linked to one parent in the upper layer and three causal neighbors in its own layer. A pixel in the root has no parent and only the causal neighbors, if any.

2.2. The proposed fusion method

The proposed method is based on three key-ideas: (i) addressing multiresolution fusion through inference on the hierarchical MMRF; (ii) fusing data collected by distinct sensors at different resolutions by inserting them in the corresponding layers of the quadtree; and (iii) fusing data acquired by different sensors at the same resolution by integrating the hierarchical MMRF with decision tree ensembles. Given a set of well-coregistered images taken by distinct VHR sensors at different resolutions on the same area, each image is included in a corresponding layer of the quadtree. Based on the quadtree, the resolutions of the images should be in a power-of-2 relation. This is a restriction, but given the resolutions of current VHR sensors, it can be easily met up to a minor resampling.

With the same notations as in Section 2.1, x_s is now meant as the class label of site $s \in S$ and a finite set Ω of classes is assumed to be defined by a training map ($x_s \in \Omega, s \in S$). As the input multisensor and multiresolution images are framed in the quadtree, each $s \in S$ is also associated with a feature vector y_s , i.e., a random field $\mathcal{Y} = \{y_s\}_{s \in S}$ of observations is defined. The proposed method is based on the marginal posterior mode (MPM) criterion, i.e., $s \in S$ is assigned to the class $\omega \in \Omega$ that maximizes $P(x_s = \omega | \mathcal{Y})$ [3]. MPM is especially appropriate for hierarchical MRFs because it penalizes errors based on the scale at which they are made – a desirable property to avoid error accumulation across the tree [4].

The proposed hierarchical MMRF is causal both spatially (see (1)) and across scales (see (2)), so an efficient recursive inference algorithm can be formulated to compute $P(x_s | \mathcal{Y})$ on each $s \in S$ [2, 4] (see Section 2.3). On one hand, this recursive process intrinsically addresses the joint classification of the data collected by sensors operating at different resolutions. On the other hand, if data are available from two sensors operating at the same resolution, then their inclusion in the quadtree yields different subsets of features on the same layer. A nonparametric ensemble approach is used to incorporate their joint distribution within the MPM and benefit from multisensor data in this case as well. Tree ensembles [7] are used for this purpose and are integrated into the aforementioned hierarchical MMRF (see Section 2.3).

2.3. MPM inference and role of tree ensembles

Based on (1) and (2) and under suitable conditional independence assumptions (analogous to those discussed in [8])

with regard to a hierarchical MRF for multitemporal classification), it can be proven that ($s \in S$) [2]:

$$P(x_s) = \sum_{x_{s-}} P(x_s|x_{s-})P(x_{s-}), \quad (3)$$

$$P(x_s|y_s^d) \propto P(x_s|y_s) \prod_{t \in s^+} \sum_{x_t} \frac{P(x_t|y_t^d)P(x_t|x_s)}{P(x_t)}, \quad (4)$$

$$P(x_s^c|x_s, y_s^d) \propto \frac{P(x_s|y_s^d)P(x_s|x_{s-})P(x_{s-})}{P(x_s)^{n_s}} \cdot \prod_{r \lesssim s} P(x_s|x_r)P(x_r), \quad (5)$$

$$P(x_s|\mathcal{Y}) = \sum_{x_s^c} P(x_s^c|x_s, y_s^d)P(x_{s-}|\mathcal{Y}) \prod_{r \lesssim s} P(x_r|\mathcal{Y}), \quad (6)$$

where y_s^d collects the feature vectors of all descendants of s in the tree, x_s^c collects the labels of all sites connected to s (i.e., x_{s-} and $\{x_r\}_{r \lesssim s}$), and n_s is the number of such sites.

Hence, MPM inference on the quadtree is accomplished through three recursive steps. First, (3) is used to recursively compute $P(x_s)$ on all sites through a top-down pass from the root to the leaves. Then, (4) and (5) are used to calculate $P(x_s^c|x_s, y_s^d)$ through a bottom-up pass from the leaves to the root. Finally, (6) is used to derive $P(x_s|\mathcal{Y})$ through a second top-down pass. Details on the initialization of these recursions and on the parametric modeling of the transition probabilities $P(x_s|x_{s-})$ and $P(x_s|x_r), r \lesssim s$, can be found in [2, 8].

Here, in particular, the symmetric mesh formulation in [6] is used to prevent anisotropic artifacts. In the case of a planar lattice, this approach is based on the definition of an appropriate visiting scheme on the pixel grid, which ensures corner independence [6]. In the proposed hierarchical model, this visiting scheme is applied within each layer of the quadtree when it is reached by the aforementioned recursive steps.

The input image data are incorporated in the recursions through the pixelwise posteriors $P(x_s|y_s)$ in (4). These posteriors are estimated through a decision tree ensemble. For this purpose, random forest [9], rotation forest [10], Extra-Trees [11], and gradient boosted regression trees (GBRT) [12] have been combined with the proposed approach. They define random collections of decision trees whose outputs are fused to predict the pixelwise posterior distribution on each site. Algorithm details can be found in [9, 10, 11, 12]. Here, we only recall that they differ in the approach taken to the bias-variance dilemma and that their applicability to highly heterogeneous features makes them especially effective for the use with multisensor imagery [13].

3. EXPERIMENTAL VALIDATION

The proposed method was experimentally validated with a VHR data set collected over Port-au-Prince, Haiti, shortly after the 2010 earthquake. It consists of an RGB GeoEye-1

image at 1.25-m resolution, an RGB-NIR QuickBird image at 2.5-m resolution, and a SAR COSMO-SkyMed stripmap image at 5-m resolution. For the COSMO-SkyMed stripmap modality, 5-m is the native resolution. The resolution of the multispectral channels of QuickBird is 2.4 m, so minor resampling was applied. In the case of GeoEye-1, the resolutions of the multispectral and panchromatic channels are 1.84 and 0.46 m, respectively. The 1.25-m image was obtained by resampling after pansharpening. The lattice size at the finest resolution is 1040×1360 pixels. Non-overlapping training and test sets were manually annotated in homogeneous areas by a specialist. No borders were taken into account in the training and test sets to avoid mixed pixels in the ground truth. Five classes are present (see Table 1).

For comparison purposes, firstly, a planar MRF classifier was used, in which unary potentials were based on the random forest predictions obtained after resampling all the images to 1.25 m. The pairwise potential was the Potts model, and the energy minimization was through sequential tree-reweighted message passing (TRW-S). Secondly, the method in [14] for the classification of multiresolution optical images through MRF and graph cuts was adapted to our multisensor data set. This method is based on Gaussian class-conditional statistics. For the SAR image, a log-normal class-conditional model was conveniently assumed and the method was extended accordingly. Details of the classification maps are shown in Fig. 1. Accuracies on the test set are in Table 1.

4. DISCUSSION AND CONCLUSION

The map obtained by the proposed approach exhibited remarkable visual regularity, thus suggesting the effectiveness of the proposed contextual technique. While the previous methods provided accurate results as well, the proposed technique improved class-wise accuracy. This confirms the effectiveness of the proposed approach to multiresolution and multisensor fusion based on the integration of hierarchical mesh and ensemble modeling, as compared both to a classical resampling approach and to an advanced multiresolution model based on linear mixture concepts [14]. Moreover, the proposed method intrinsically generates maps at all resolutions (an often desirable property when multiresolution data are available); the previous methods only classify at one resolution.

The accuracies of the proposed method when paired with the above mentioned tree ensembles were rather similar at the 1.25 and 2.5-m resolutions, the highest accuracies being achieved using random forest and GBRT. However, a substantial improvement ($\geq 6\%$) was obtained using GBRT over the other considered ensembles at the resolution of 5 m. On one hand, these results confirm the flexibility of the proposed approach in incorporating predictions from arbitrary ensemble methods. On the other hand, they suggest GBRT as an especially appropriate choice within the proposed framework.

Table 1. Test-set accuracies of the proposed method, applied with random forest (RanFor), rotation forest (RotFor), ExtraTrees, and GBRT, and of the compared benchmark methods.

resolution 1.25 m	container %	vegetation %	asphalt %	buildings %	sea %	over all %	Cohen's kappa coeff
Proposed method, RanFor	85.76	34.29	95.03	98.86	95.35	96.06	0.9047
Proposed method, RotFor	83.37	38.75	81.51	99.52	67.73	94.48	0.8613
Proposed method, ExtraTrees	82.29	37.63	88.75	98.77	91.89	95.07	0.8786
Proposed method, GBRT	87.08	33.27	95.17	99.04	97.18	96.36	0.9124
Single-res. MRF after resampling	63.97	74.81	98.64	99.30	76.24	94.58	0.888
Adaptation of [14]	75.06	85.32	95.89	97.75	99.36	95.11	0.901

resolution 2.5 m	container %	vegetation %	asphalt %	buildings %	sea %	over all %	Cohen's kappa coeff
Proposed method, RanFor	85.95	27.51	95.32	99.58	95.85	96.23	0.9174
Proposed method, RotFor	76.85	32.71	81.18	100.00	92.29	94.47	0.8691
Proposed method, ExtraTrees	84.76	32.71	95.77	99.44	95.45	96.13	0.9148
Proposed method, GBRT	86.31	32.71	94.87	100.00	96.04	96.63	0.9276

resolution 5 m	container %	vegetation %	asphalt %	buildings %	sea %	over all %	Cohen's kappa coeff
Proposed method, RanFor	73.32	15.87	84.40	97.28	32.54	90.00	0.7680
Proposed method, RotFor	45.91	14.28	65.60	97.26	28.57	85.81	0.6561
Proposed method, ExtraTrees	57.93	20.63	72.24	97.23	31.74	87.58	0.7035
Proposed method, GBRT	87.98	25.39	96.55	100.00	88.88	96.01	0.9288

5. REFERENCES

- [1] I. Hedhli, G. Moser, S. B. Serpico, and J. Zerubia, "Classification of multisensor and multiresolution remote sensing images through hierarchical Markov random fields," *IEEE Geosci. Rem. Sens. Lett.*, vol. 14, no. 12, pp. 2448–2452, 2017.
- [2] I. Hedhli, G. Moser, S. B. Serpico, and J. Zerubia, "Multi-resolution classification of urban areas using hierarchical symmetric Markov mesh models," in *Proc. of JURSE 2017*, 2017.
- [3] S.Z. Li, *Markov Random Field Modeling in Image Analysis*, Springer, 2009.
- [4] J. Laferté, P. Pérez, and F. Heitz, "Discrete Markov image modeling and inference on the quadtree," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 390–404, 2000.
- [5] P. A. Devijver, "Hidden Markov mesh random field models in image analysis," *J. Applied Stat.*, vol. 20, no. 5-6, pp. 187–227, 1993.
- [6] S. Yousefi, N. Kehtarnavaz, and Y. Cao, "Computationally tractable stochastic image modeling based on symmetric Markov mesh random fields," *IEEE Trans. Image Process.*, vol. 22, no. 6, 2013.
- [7] A. Criminisi and J. Shotton, *Decision Forests for Computer Vision and Medical Image Analysis*, Springer, 2013.
- [8] I. Hedhli, G. Moser, J. Zerubia, and S. B. Serpico, "A new cascade model for the hierarchical joint classification of multitemporal and multiresolution remote sensing data," *IEEE Trans. Geosci. Rem. Sens.*, vol. 54, no. 11, pp. 6333–6348, 2016.
- [9] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [10] J. J. Rodriguez, L. I. Kuncheva, and C. J. Alonso, "Rotation forest: A new classifier ensemble method," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1619–1630, 2006.
- [11] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 63, no. 1, pp. 3–42, 2006.
- [12] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Stat.*, vol. 29, no. 5, pp. 1189–1232, 2001.
- [13] A. Merentitis and C. Debes, "Many hands make light work - on ensemble learning techniques for data fusion in remote sensing," *IEEE Geosci. Rem. Sens. Mag.*, vol. 3, no. 3, pp. 86–99, 2015.
- [14] G. Moser, A. De Giorgi, and S. B. Serpico, "Multiresolution supervised classification of panchromatic and multispectral images by Markov random fields and graph cuts," *IEEE Trans. Geosci. Rem. Sens.*, vol. 54, no. 9, pp. 5054–5070, 2016.

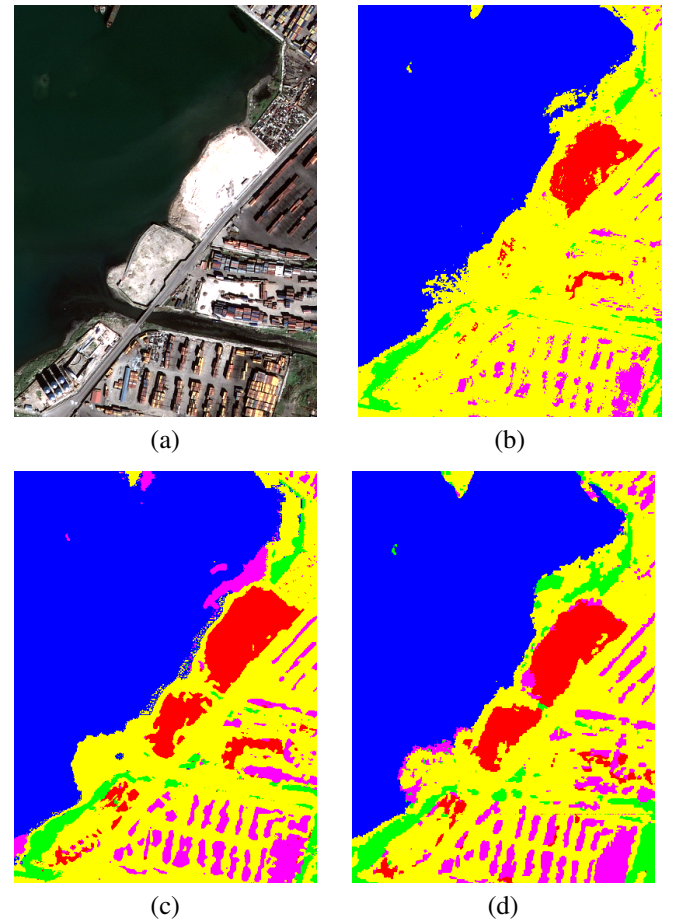


Fig. 1. (a) Details of the optical image at 1.25-m resolution and of the classification maps obtained using: (b) single-res. MRF after resampling, (c) adaptation of [14], and (d) the proposed method with GBRT.